mtDNA diversity in human populations highlights the merit of haplotype matching in gene therapies

Running title: **Implications of mtDNA diversity for gene therapies**

E.C. Røyrvik[1], J.P. Burgstaller[2,3], I.G. Johnston[4]*

[1]Division of Biomedical Sciences, Warwick Medical School, University of Warwick, Coventry CV4 7AL, UK, [2]Biotechnology in Animal Production, Department for Agrobiotechnology, IFA Tulln, 3430 Tulln, Austria, Institute of Animal Breeding and Genetics, University of Veterinary Medicine, Vienna, Veterinärplatz 1, 1210 Vienna, Austria, [4]School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

Abstract

Modern gene therapies aim to prevent the inheritance of mutant mitochondrial DNA (mtDNA) from mother to offspring by using a third-party mtDNA background. Technological limitations mean that these therapies may result in a small amount of maternal mtDNA admixed with a majority of third-party mtDNA. This situation is unstable if the mother's mtDNA experiences a proliferative advantage over the third-party mtDNA, in which case the efficacy of the therapy may be undermined.

Animal models suggest that the likelihood of such a proliferative advantage increases with increasing genetic distance between mother and third-party mtDNA, but in real therapeutic contexts the genetic distance, and so the importance of this effect, remains unclear. Here we harness a large volume of available human mtDNA data to model random sampling of mother and third-party mtDNAs from real human populations. We show that even within the same haplogroup, genetic differences around 20-80 SNPs are common between mtDNAs. These values are sufficient to lead to substantial segregation in murine models, over an organismal lifetime, even given low starting heteroplasmy, inducing increases from 5% to 35% over one year. Randomly pairing mothers and third-party women in clinical contexts thus runs the risk that substantial mtDNA segregation will compromise the beneficial effects of the therapy. We suggest that choices of 'mtDNA donors' be based on recent shared maternal ancestry, or, preferentially, explicit haplotype matching, in order to reduce the potential for problems in the implementation of these therapies.

Introduction

Mitochondria are small organelles within eukaryotic cells that are vital for the normal aerobic production of ATP, the 'universal' biochemical energy carrier. Each mitochondrion, of which there are many in any given cell, carries at least one copy of its own, small genome (mitochondrial or mtDNA), distinct from the large genome stored in the nucleus. While there are good reasons for retaining some genes in

the mitochondrion (Johnston and Williams, 2016), a challenging biochemical environment and comparative lack of efficient DNA repair mechanisms allows a higher mutation rate there than in the nucleus (Alexeyev *et al.* , 2013).

Differences in the sequence of mitochondrial DNA can arise at the level of individuals (population diversity) or different mitochondria in the same cell (*heteroplasmy* – see below). In humans, mtDNA is inherited uniparentally, via the mother's egg cell; recombination is usually negligible between human mtDNAs (Hagelberg, 2003, Hagstrom *et al.* , 2014). Given the non-recombining nature of the mitochondrial genome, such polymorphisms as exist can be expressed in terms of a straightforward phylogenetic tree (see Fig. 1A). The sum of polymorphisms in an mtDNA sequence is known as a haplotype, and any hierarchical clade of haplotypes is a haplogroup. Since inheritance is uniparental, mtDNA haplogroups are strongly susceptible to genetic drift, and this has given rise to pronounced haplogroup pattern differences between geographical areas, especially on a continental scale (see Fig. 1B).

Mitochondrial diversity in humans is often neutral or near-neutral (Chinnery and Hudson, 2013), although an increasing volume of research in animal models and humans suggests that non-pathogenic mtDNA variants can be associated with some phenotypic effects, from livestock fertility to longevity and disease

susceptibility (Dowling, 2014, Latorre-Pellicer *et al.* , 2016, St John, 2016, Tsai and St John, 2016, Wallace, 2015, Wallace and Chalkia, 2013). We note that, while evidence exists for a range of phenotypic effects, flawed analyses have in some cases led to several statistically unsupported claims of mtDNA links to disease (Johnston, 2015).

While the phenotypic effects of some mtDNA variants are relatively mild, certain mtDNA mutations in humans have dramatic phenotypic consequences, causing fatal, incurable diseases (for example, mt3243A>G, causing the inherited disease MELAS), which often manifest when the proportion of mutated mtDNA molecules in a cellular population exceeds a threshold (Taylor and Turnbull, 2005, Wallace and Chalkia, 2013). Clinical approaches to prevent the inheritance of diseases resulting from damaging mutations in mtDNA are a focus of current medical research. Cutting-edge therapies including pronuclear transfer and chromosomal spindle transfer attempt to address the inheritance of mutant mtDNA from a maternal carrier by transferring the nuclear genome (either as the pair of pronuclei or the chromosomal spindle) into a third-party, enucleated oocyte or zygote with non-pathogenic mtDNA (Brown *et al.* , 2006, Burgstaller *et al.* , 2015, Craven *et al.* , 2010, Tachibana *et al.* , 2009) (Fig. 2). These therapies thus aim to place parental nuclear DNA on a healthy mitochondrial background with no mtDNA from the mother present. However, technological limitations currently mean that *carryover* is possible, whereby some of the mother's mtDNA may be carried into the third-party

cell with the transferred nuclear genetic material. These therapies can thus lead to the coexistence of several distinct sequences within cellular mtDNA populations. First, the non-pathogenic mtDNA from the third-party oocyte donor is present. Second, due to carryover, non-pathogenic mtDNA from the mother may be present. Third, due to carryover, pathogenic (mutant) mtDNA from the mother may be present (Fig. 2). The resulting complex system may give rise to phenotypic effects due to differences between admixed mtDNA types (Burgstaller *et al.*, 2015) and references therein, and between the nucleus and different mtDNA types (reviewed in (Reinhardt *et al.* , 2013)), highlighted by very recent work in mouse model (Latorre-Pellicer *et al.*, 2016). Previous work has reviewed the potential implications of these effects on gene therapies (Morrow *et al.* , 2015, Reinhardt *et al.*, 2013). In this article we will focus on the possibility, and implications, of proliferative differences between different mtDNA types.

The above admixture of mtDNA types is stable if mother and oocyte donor mtDNA experience no proliferative differences (Fig. 2, centre), and if the oocyte donor haplotype experiences a proliferative advantage then carried-over mtDNA will generally be reduced over time (Fig. 2 left). However, a general proliferative advantage of the mother's haplotype can in principle lead to the amplification of the associated pathological mutation, working against the desired effect of the therapy to remove this mutation (Fig. 2 right). *This amplification can in principle occur even if the pathological mutation itself experiences a selective disadvantage* – if this

disadvantage is of lower magnitude than the proliferative difference between haplotypes, the latter effect will still dominate.

In a wide selection of mammalian species, such proliferative differences between mtDNA haplotypes have been shown to exist (St John *et al.* , 2010). Pronounced differences have been shown in various mouse models e.g. (Burgstaller *et al.* , 2014, Sharpley *et al.* , 2012), pigs (Takeda *et al.* , 2006), mini-pigs (Cagnone *et al.* , 2016), and cattle (Ferreira *et al.* , 2010). Sets of models and studies exhibiting this behaviour are reviewed in (Burgstaller *et al.*, 2015, St John, 2012). Recent work in human cell lines (Yamada *et al.* , 2016) has illustrated that pronounced changes in the balance of mtDNA haplotypes in cellular populations can occur over time, with an initially small population of one haplotype (H1) becoming dramatically amplified and subsequently reduced through cell passages when admixed with a distantly-related human haplotype (L3). Recent results from a human stem cell line ultimately derived from an instance of pronuclear transfer explicitly demonstrate that amplification of carried-over mtDNA can occur after therapy implementations, in some instances from 4% to >40% of the cellular population over 10 passages, even with genetically similar (same haplogroup) mtDNA sequences (Hyslop *et al.* , 2016).

While the direction and tissue-dependence of differential proliferation are currently difficult to predict for a given system, the expected magnitude of the difference depends on the genetic distance between haplotypes (Burgstaller *et al.*, 2014) (Fig.

3). An important question to consider in gene therapies is thus, given the mtDNA diversity in human populations, what genetic distances are likely to arise in nuclear mother-oocyte donor pairings in therapeutic contexts, and what is the magnitude of the proliferative differences (Fig. 2) these distances will produce?

If $\prod_{ij}$ is the number of non-identical bases between two mtDNA genomes, *i* and *j*, then, intuitively, identical mtDNAs ($\prod_{ij} = 0$) would be expected to behave identically, but the more different the mtDNAs ($\prod_{ij} > 0$), the larger is the proliferative difference generally expected between the two. We define *heteroplasmy*, *h*, as the proportion of one 'foreign' mtDNA haplotype in a cellular admixture: hence, if a cell contains H_0 mtDNAs of its 'native' haplotype and H_1 mtDNAs of a 'foreign' haplotype, $h = H_1/(H_0 + H_1)$.

Proliferative differences between haplotypes can be measured as a quantity β, a rate of proliferation of one mtDNA over another, overcoming the limitations inherent in considering absolute differences in heteroplasmy percentages (see SI for a formal definition). For example, proliferative differences of average magnitude $|\beta| \simeq 0.008$ per day have been measured between two mtDNA types of $\prod_{ij} \simeq 100$ in the livers of mice; this value of β corresponds to an amplification of *h* from 0.05 (5% of one haplotype) to 0.49 (49% of that same haplotype) over one year (Burgstaller *et al.*, 2014). This pronounced rate of change is supported by results in

a range of other mammalian models (including rapid fixation of an initial limited mtDNA haplotype in cattle (Koehler *et al.* , 1991) and the aforementioned results from human cell lines where changes from <10% to >40% occur over a small number of cell passages (Hyslop *et al.*, 2016).

A subset of recent evidence for proliferative differences between mtDNA haplotypes in mice is shown in Fig. 3. Fig. 3A shows inferred values of $|\beta|$, and the magnitude of proliferative differences between mtDNAs, in a variety of tissues for three mtDNA pairs (where $\prod_{ij} = 18$, 86, and 107). Fig. 3B shows the predictions that this behaviour of β makes about absolute changes in heteroplasmy, for two putative admixtures beginning with 5% and 20% of a 'foreign' haplotype. For example, a haplotype differing from the 'native' type by $\prod_{ij} \simeq 100$ may readily experience amplification from 5% to 50% over one year.

For simplicity, these plots are limited to the behaviour over one year, but the trends are observed to continue throughout organismal lifetimes. For example, one observation in (Burgstaller *et al.*, 2014) showed heteroplasmy in liver tissue rising from 5.9% to 81.8% over 680 days for a particular mtDNA pairing where $\prod_{ij} = 108$. There is thus evidence that, in mice, nucleotide differences around $\prod_{ij} \sim 100$ are associated with proliferative differences capable of amplifying an admixed haplotype from a 5% minority to a pronounced cellular majority over the course of an organismal lifetime. But what are standard values of $\prod_{ij}$ in actual human populations? And is this magnitude of genetic diversity expected to give rise to clinically relevant mtDNA behaviour, given that a mutant mtDNA load of 40-60% is

often sufficient to cause morbidity, and it still poorly known what ‘safe’ levels may be in most cases (Wallace and Chalkia, 2013)?

Existing studies have characterised the nucleotide differences in contemporary human populations, finding typical differences of dozens of nucleotides across modern Europeans (Fu *et al.* , 2012), greater diversity in Africa than in Europe (Briggs *et al.* , 2009), and results confirming and expanding these observations across a broader geographical range (Lippold *et al.* , 2014). A modern workflow has been developed to address related evolutionary questions (Blanco *et al.* , 2011). However, to our knowledge, the interpretation of these statistics in terms of mtDNA segregation possibility and implications for disease is currently absent, as is an attempt to characterise the expected diversity in modern populations combining social (census) and biological (sequence) data.

Materials and Methods

Materials – None.

Methods – We took a data-driven approach, harnessing the large numbers of human mtDNA sequence data now available through the NCBI database, as well as haplogroup data in the literature. mtDNA molecules may be categorised, via the presence or absence of diagnostic SNPs, into haplogroups, which are typically designated by an alphanumeric code and follow a moderately complex hierarchy. For example, at the coarsest level, all human mtDNAs so far recorded fall into haplogroup *L*. Subsets of *L* include *N* (which in turn includes *R*, containing *H* and

194 *V, etc.*) and *W, X, Y* and others. A simplified tree of haplogroups is shown in Fig.
195 1A and illustrative geographical distributions are shown in Fig. 1B.
196 Data on the haplogroup makeup of 'pre-colonial populations', i.e. before early
197 modern population mixing, from different geographical regions is available via
198 MitoMAP (Lott *et al.* , 2013). These data can be used to estimate the probability
199 that an individual with maternal ancestry from a given region belongs to a given
200 haplogroup.
201 Many specific mtDNA sequences corresponding to individual humans belonging to
202 a given haplogroup are available via NCBI. Using these data, we sought to identify
203 the expected genetic differences between pairs of individual, real human mtDNAs.
204 To estimate these expected differences, we first characterised the expected
205 differences between specific mtDNA samples within and between different
206 haplogroups.
207 We obtained the $> 30k$ mtDNA sequences available from NCBI Nucleotide
208 database (NCBI, 2015). Of these sequences $\sim 7.6k$ had straightforwardly
209 interpretable haplogroup information, where the initial letter of the */haplogroup* field
210 was taken to be the haplogroup label. We categorised these records by this initial
211 letter, then employed the following sampling protocol. Given a pair of haplogroups
212 $\{\mathcal{H}_1, \mathcal{H}_2\}$, we picked at random a sequence belonging to $\mathcal{H}_1$ and picked at random
213 a sequence belonging to $\mathcal{H}_2$ (ensuring that the two sequences were not the same
214 sample if $\mathcal{H}_1 = \mathcal{H}_2$,). We used BLAST to record the number of sequence
215 differences between these specific sampled sequences. For the purposes of this

report we recorded the number of non-identical bases as the nucleotide difference $\prod_{ij}$; we also note that indels commonly exist between sampled mtDNA sequences, further contributing to mtDNA diversity. We then built up a distribution of sequence differences over many (*n=1000*) sampled pairs of specific human mtDNAs from the given pair of haplogroups.

To connect more explicitly with medical policy, we next changed the scale of our analysis from haplogroups *per se* to the estimated haplogroup profiles of real human populations. First, we employed heuristic data from the MitoMAP project (Lott *et al.*, 2013) estimating the haplogroup makeup of pre-colonial populations from different regions of the world, while noting that the actual census populations will usually have a very different makeup, especially in New World countries that experienced extensive overseas colonization. For each region, we randomly chose two haplogroups, each with a probability corresponding to that haplogroup's representation in the region of interest. We then randomly chose two specific mtDNA sequences from those two haplogroups. As above, we then used BLAST to determine the genetic difference between those specific sequences. We repeated this process many times to build up an expected distribution of the genetic differences between two randomly chosen members of the human population from that region.

As the UK is on the cusp of implementing gene therapies based on nuclear transfer, we then performed a more rigorous, population-based analysis for Britain. In order to estimate the probable levels of nucleotide diversity ($\prod_{ij}$) in mtDNA

238 between two randomly selected British women, and hence the likely magnitude of
239 proliferative differences between their mtDNA, a haplogroup profile of Britain was
240 assembled, based on over 4,600 individuals. The majority of the UK samples
241 represent ethnic Britons. To account for the fact that the modern UK population
242 consists of many ethnicities, approximations of mtDNA haplogroup distributions for
243 the two largest cities in the UK (London and Birmingham) were also constructed.
244 These distributions are estimates, based on data from the 2011 census,
245 immigration data, and published mtDNA haplogroup data for areas from which
246 there has been mass immigration into the UK (see SI for details).
247 For each ethnic census category, an estimate of probable haplogroup composition
248 was created (see SI for details on calculations), and the frequency values scaled
249 by the numerical census data to yield expected haplogroup frequencies in London
250 and Birmingham. For simplicity, the single letter level of nomenclature is used,
251 with the exception of superhaplogroup *L*, for which its subgroups *L0-3* are included.
252

253 **Results**

254 Fig. 4A shows the resulting statistics on differences between sampled mtDNA
255 sequences between haplogroup pairs. Several intuitive features are immediately
256 observable. First, haplogroup *L* displays noticeably more intra-haplogroup
257 differences than any other haplogroup. *L* haplogroups constitute the majority of
258 African haplogroups (and have very deep branching times relative to non-African
259 haplogroups) and are thus expected to include the most genetic diversity (Behar *et*

al. , 2008). Second, with the exception of *L*, diagonal elements (i.e. samples from a haplogroup compared to samples from the same haplogroup) show less diversity than off-diagonal elements (i.e. samples from one haplogroup compared to samples from a different haplogroup). Third, haplogroup pairings which are expected to be similar (for example, sister clades *H* and *V*) show decreased genetic diversity. The inset shows a breakdown of the *L* haplogroup into its immediate subgroups.

A notable result from this analysis is that between haplogroups, differences of ~50 SNPs are common, and, even within haplogroups, differences of ~20 SNPs are not uncommon. This level of diversity may not seem substantial when compared to the ~16 kilobases of total human mtDNA, but we draw attention to our previous observations that differences of ~20 SNPs were enough to induce significant proliferative differences between haplotypes in mice, who also have a ~16kb mtDNA genome (Burgstaller *et al.*, 2014). As shown in parentheses in Fig. 4A, the magnitudes of $\prod$ that likely emerge from pairwise haplotype samples match those responsible for dramatic mtDNA heteroplasmy changes in mouse models.

Fig. 4A also provides a means of identifying a ‘partner’ for a given haplogroup that minimizes $\prod$ and hence the likelihood of damaging segregation. For example, given a mother with haplogroup B and a choice between donors from C, V, and L, Fig.4A shows that the B-V pairing minimizes maximum $\prod$, and thus affords the lowest risk of high segregation (see Discussion).

Table 1 gives the estimated haplogroup makeup of the UK and two major cities, based on a combination of census and immigration data and a survey of worldwide mtDNA sequences (see Methods and SI). We underline that these quantities are principled estimates, but the summary statistics that arise from these estimates are robust to variation in the exact population frequencies, and is consistent with the behaviour expected from an ethnically mixed population based on more direct estimates (see below).

Fig. 4B illustrates the distribution of nucleotide differences between individuals sampled from geographical regions, and rural vs. urban UK based on estimates in Table 1, in this manner. It is immediately noticeable that pairs of individuals from Africa generally exhibit more diversity than pairs chosen from other regions, but it is striking that the expected genetic difference in many geographic regions is around $\prod_{ij} \sim$ 40-50 SNPs, often with a range between 10-100 SNPs. The substantial diversity expected in the UK and its cities is of a consistent magnitude with that expected from its population history, involving admixtures of African and Asian immigrants in addition to its original European state. Again, parenthesized numbers in Fig. 4B illustrate that these magnitudes of $\prod$ are readily able to induce pronounced heteroplasmy shifts in mice. Taken together, these results demonstrate that expected levels of mtDNA diversity in modern human populations are of comparable magnitude to those responsible for substantial segregation bias in existing mammalian models, and so therapies that randomly pair women from

these populations may engender potentially detrimental heteroplasmy changes over time.

Discussion

Our analysis clearly shows that, even within a geographical region restricted to the point of being dominated by a single mtDNA haplogroup, a $\prod_{ij} = 10 - 100$ is expected between randomly sampled individuals from that region. On a continental scale, expected differences are highest in Africa, as predicted from our knowledge of human population history, and comparably lower elsewhere. Comparably high, however, are the differences in the largest urban populations of the UK, where oocyte donor therapies will be implemented.

In mice, proliferative differences between haplogroups with $\prod_{ij} \sim 100$ were sufficient in some tissues to cause amplification of one mtDNA type from 0.05 to 0.64 (i.e. a small representation to a notable majority) over an organismal lifetime (Fig. 2B). There remains a wide range of questions involving the mapping from the murine model to the human system. One criticism of our argument may be that mtDNA segregation in humans may progress more slowly than in mice, reducing the magnitude of the effects we consider. However, segregation in humans has been observed to occur more rapidly than in mice (Wallace and Chalkia, 2013). Furthermore, evidence exists for pronounced segregation of a pathological mutation over very short times during embryo-fetal development (Monnot *et al.* , 2011), suggesting the presence of mechanisms in humans that support fast

segregation, and which could in principle also act on non-pathological mutations. Recent results in human cell lines (Hyslop *et al.*, 2016, Yamada *et al.*, 2016) showing fast changes in mtDNA population structure over passages support the possibility of fast segregation. These rapid mtDNA dynamics are supported by evidence from other large mammalian models, including the rapid fixation of mtDNA haplotypes in cattle (Burgstaller *et al.*, 2015, Koehler *et al.*, 1991) Even in a conservative case where mtDNA turnover rates are scaled by organismal lifetimes, amplification over the (longer) human lifetime will still be anticipated by analogy with the murine system. An important clinical example of the potentially high mtDNA segregation in human disease (again involving a pathological mutation) is described in Ref.(Mitalipov *et al.* , 2014), in which an embryo selected for its low (12%) load of the 3243 mutation (Treff *et al.* , 2012) developed into an infant with >40% loads in blood and urine at six weeks of age, presenting with a range of (possibly unrelated) metabolic pathologies.

It is worth noting that, in addition to the unpredictability of segregation direction, the rate at which mtDNA segregation occurs is not simple and constant – rather, it can depend on tissue type, organismal age and developmental stage (Burgstaller *et al.*, 2014), and complicating processes including the mtDNA bottleneck (Johnston *et al.* , 2015). In addition, increasing evidence that mtDNA variants may influence fertility and development (St John, 2012, St John *et al.*, 2010) suggests further potential complications as mtDNA populations both influence and are influenced by developmental dynamics. Given these complications, it is not unreasonable to think

346 that the ‘averaged’ rates reported here may be underestimates for a particular time
347 period. We therefore highlight that, even from a conservative calculation of
348 segregation rates, *the likely genetic differences between humans randomly*
349 *sampled from a population may well allow substantial amplification of a*
350 *disease-carrying mtDNA haplotype over the timescale of a human lifetime*.
351
352 We must also consider whether randomly sampling NCBI sequences is a good
353 model for the mtDNA pairings likely to be involved in gene therapies. The counter-
354 example of this would be a population consisting of many individuals with identical
355 mtDNA sequences and a small number of individuals with different sequences. The
356 NCBI, which assigns records to unique sequences, will likely have one record for
357 the common sequence and one each for the rare different sequences. In this case,
358 uniformly sampling NCBI would underestimate the population fraction with the
359 common sequence, and thus tend to overestimate mtDNA diversity. However, the
360 ubiquity of many-SNP differences between records (see Fig.4) suggests that this
361 problematic population structure is unlikely, and indeed, several contemporary
362 studies have observed differences between each individual sample (Fu *et al.*,
363 2012, Lippold *et al.*, 2014). Additionally, socio-economic factors will give rise to
364 structure in the pairings in clinical applications (which may either decrease or
365 increase the expected $\prod_{ij}$). Despite these complications, we consider our
366 approximations appropriate for considering first-order bounds of likely behaviour in
367 these populations exhibiting realistic human diversity.

The danger of pathological mutations ‘hitchhiking’ on favoured haplotype backgrounds and being amplified along with the haplotype is described in the introduction and has been discussed previously (Burgstaller *et al.*, 2014, Burgstaller *et al.*, 2015). An additional danger is the amplification of an initially rare mtDNA haplotype to the point where it competes with the dominant mtDNA type in a cell and causes pathologies through mismatched mitochondrially encoded protein subunits or other mechanisms (Burgstaller *et al.*, 2015). The co-occurrence in a cell of two different, but both separately non-pathogenic, mtDNAs has been observed to result in adverse physiological changes (Sharpley *et al.*, 2012), and so-called mito-nuclear incompatibilities between nuclear and 'foreign' mtDNA content can induce phenotypic effects (Latorre-Pellicer *et al.*, 2016) – resulting in potential implications for gene therapies that have been reviewed elsewhere (Morrow *et al.*, 2015, Reinhardt *et al.*, 2013). Segregation between mtDNA haplotypes, allowing an initially rare haplotype to proliferate and become amplified within a cell, has the potential to manifest and exacerbate all of these potential issues.

To diminish the likelihood of potentially harmful mtDNA segregation, which we argue is likely given the mtDNA diversity in the modern UK population, we urge experts involved in the implementation of these therapies to consider ‘haplotype matching’, i.e. choosing an oocyte donor with mtDNA as similar as possible to the mother's in clinical approaches. Methods to match haplotypes (minimise $\prod_{ij}$) could

include choosing maternal relatives of the mother with low or zero proportions of the pathological mutation under consideration, or choosing donors from a haplogroup as similar as possible to the mother's. To illustrate this latter strategy, Fig. 5 shows the range of expected $\prod$ values that could arise when a third-party donor is paired with a mother from haplogroup H1a. If no haplotype matching is employed, and the third-party donor is randomly sampled from our estimated London population, a maximum $\prod$ around 100 is possible (due to the pronounced population diversity illustrated in Fig. 4B). Choosing a donor from haplogroup H decreases this maximum value to around 36 (that is, the maximal within-H diversity, shown on the diagonal of Fig. 4A). More detailed matching, specifically choosing another H1a woman as the third-party, further limits the maximum $\prod$ to approximately 17. These lower values achieved through haplotype matching dramatically decrease the expected potential heteroplasmy changes (for example, in mice (Fig 2), from a maximum of 5% $\rightarrow$ 49% over one year for $\prod = 100$ to 5% $\rightarrow$ 8% over one year for $\prod = 17$), thus immediately limiting the potential for detrimental segregation. Our results, and future findings from more detailed studies, can help provide a strategy for this matching process – given a mother of known mtDNA haplogroup, choose from available oocyte donors so as to minimise the maximum genetic distance given in Fig. 4. Such haplotype matching, which is in principle technically straightforward and economically marginal, decreases the risk of inadvertently choosing an mtDNA pairing which experiences substantial

proliferative differences, and thus decreases the risk of manifestation of the disease the therapy was implemented to prevent.

Table 1. Estimated haplogroup frequencies in the British population UK – majority ethnic Britons, exclusive of large urban areas, London, Birmingham – census and immigration data based estimates (see SI).

HG	UK %	London %	Birmingham %
A	0.0%	0.6%	0.5%
B	0.0%	1.1%	0.7%
C	0.0%	0.3%	0.2%
D	0.0%	0.8%	0.5%
F	0.0%	1.1%	0.8%
G	0.0%	0.2%	0.2%
H	45.2%	30.4%	29.9%
I	4.1%	2.6%	2.6%
J	12.4%	7.8%	8.2%
K	8.3%	5.1%	5.3%
L0	0.0%	1.3%	0.8%
L1	0.0%	2.4%	1.9%
L2	0.0%	4.9%	3.8%
L3	0.1%	4.5%	3.5%
M	0.0%	10.4%	12.7%
N	0.0%	0.1%	0.2%
O	0.0%	0.0%	0.0%
P	0.0%	0.0%	0.0%
R	0.1%	2.7%	3.6%
S	0.0%	0.0%	0.0%
T	10.5%	6.8%	6.9%
U	12.6%	11.5%	12.6%
V	3.2%	1.6%	1.8%
W	1.5%	1.2%	1.6%

X	1.8%	1.2%	1.1%
other	0.3%	1.3%	0.6%

Figure 1. A) Relationship between human mtDNA haplogroups. Haplogroup labels and tree structure for human mtDNA groups; *MRCA* is most recent common ancestor. **B) Typical haplogroups in pre-colonial human populations by approximate geography.** We have omitted higher-order haplogroups of which many sub-groups are presented (e.g. *N* & *R*). Based on data from MitoMAP (Lott *et al.*, 2013) and references therein.

Figure 2. mtDNA segregation and gene therapies. A mother may possess two similar haplotypes, one wild type (blue) and one mutant (blue with red star). Therapies attempt to use a third-party with a potentially different mtDNA haplotype (yellow) to provide a healthy mtDNA background. Carryover in these therapies may result in an admixture of wildtype mother, mutant mother, and wildtype third-party mtDNA in a cell. If the two haplotypes (blue and yellow) proliferate differently, the offspring may evolve a predominance of third-party (lower left) or mother (lower right) mtDNA with time. In the latter case, if mutated mtDNA proliferates at a similar rate to its ‘carrier’ haplotype, the damaging mutation may be amplified to harmful levels in cells.

Figure 3. mtDNA segregation and genetic differences in mice. A) Magnitudes of segregation (proliferative differences between mtDNA types) in different tissues (points) in four different mtDNA pairings from (Burgstaller *et al.*, 2014). More pronounced segregation is observed in those pairings with the greatest genetic distance. Red line shows the mean trend of segregation with number of nucleotide differences; blue line shows the approximate maximum segregation strength across all tissues for mtDNA pairings with < 100 nucleotide differences. **B)** Ranges of expected heteroplasmy in mice after 1 year, given different initial heteroplasmies (h_0) and the mean (lower) and maximal (higher) segregation magnitude observed in mice. For example, the darker red curve shows that for an mtDNA pairing with 75 nucleotide differences, a maximal increase from $h = 0.05$ to $h \simeq 0.3$ is expected.

Figure 4. A) MtDNA differences between haplogroups. The maximum (outer halo) and minimum (inner halo) nucleotide differences expected between a pair of randomly sampled mtDNA sequences (horizontal and vertical axes). The diagonal corresponds to pairs within the same haplogroup; off-diagonal elements correspond to pairs of mtDNAs from different haplogroups. Dataset size for each haplogroup is given in brackets; n=1000 samples were used for each pairing. Max *h* change shows, for a given magnitude of genetic diversity, the maximum expected change in heteroplasmy over one year starting at 5%, based on mouse models (Fig 3). As described in the text, haplotype labels denote sequences that fall within a given category and not within any named subcategories of that

458 category. Inset shows subgroups of the most-diverse *L* haplogroup. Red circles
459 give the magnitudes of genetic differences between the “background” C57BL/6N
460 mtDNA and the different mtDNA types in the mouse models in Fig. 3. **B) MtDNA**
461 **differences between geographical regions.** In blue, genetic differences between
462 a pair of individuals randomly sampled from sets modelling populations within a
463 given region of the world, using the MitoMAP (Lott *et al.*, 2013) estimation of the
464 (pre-colonial) haplogroup profile of different geographical regions. In black,
465 expected differences in the general the modern non-urban UK population, and
466 populations of London and Birmingham. Candlesticks show minimum, mean $\pm$ s.d.,
467 and maximum nucleotide differences between simulated pairs sampled from
468 geographical regions. Explicit sample distributions are given in in lighter colours;
469 max h change gives maximum expected change in heteroplasmy as in (A). SE
470 Asia (in grey) has poorly characterised MitoMAP estimates. Red marks, as in (A),
471 give the magnitudes of genetic differences in the mouse models in Fig. 3.
472
473 **Figure 5. MtDNA differences expected with different haplotype matching**
474 **strategies for a mother with haplogroup H1a.** Distributions of nucleotide
475 differences (min, mean +- sd, max) expected when pairing mtDNA from haplogroup
476 H1a with randomly sampled mtDNA from our estimated London population, with
477 randomly sampled mtDNA from haplogroup H, and with randomly sampled mtDNA
478 from haplogroup H1a.

Alexeyev M, Shokolenko I, Wilson G, LeDoux S (2013) The maintenance of mitochondrial DNA integrity--critical analysis and update. Cold Spring Harbor perspectives in biology 5, a012641.

Behar DM, Villems R, Soodyall H, Blue-Smith J, Pereira L, Metspalu E, Scozzari R, Makkan H, Tzur S, Comas D *et al.* (2008) The dawn of human matrilineal diversity. American journal of human genetics 82, 1130-40.

Blanco R, Mayordomo E, Montoya J, Ruiz-Pesini E (2011) Rebooting the human mitochondrial phylogeny: an automated and scalable methodology with expert knowledge. BMC bioinformatics 12, 174.

Briggs AW, Good JM, Green RE, Krause J, Maricic T, Stenzel U, Lalueza-Fox C, Rudan P, Brajkovic D, Kucan Z *et al.* (2009) Targeted retrieval and analysis of five Neandertal mtDNA genomes. Science 325, 318-21.

Brown DT, Herbert M, Lamb VK, Chinnery PF, Taylor RW, Lightowlers RN, Craven L, Cree L, Gardner JL, Turnbull DM (2006) Transmission of mitochondrial DNA disorders: possibilities for the future. Lancet 368, 87-9.

Burgstaller JP, Johnston IG, Jones NS, Albrechtova J, Kolbe T, Vogl C, Futschik A, Mayrhofer C, Klein D, Sabitzer S *et al.* (2014) MtDNA segregation in heteroplasmic tissues is common in vivo and modulated by haplotype differences and developmental stage. Cell reports 7, 2031-41.

Burgstaller JP, Johnston IG, Poulton J (2015) Mitochondrial DNA disease and developmental implications for reproductive strategies. Molecular human reproduction 21, 11-22.

Cagnone G, Tsai TS, Srirattana K, Rossello F, Powell DR, Rohrer G, Cree L, Trounce IA, St John JC (2016) Segregation of Naturally Occurring Mitochondrial DNA Variants in a Mini-Pig Model. Genetics 202, 931-44.

Chinnery PF, Hudson G (2013) Mitochondrial genetics. British medical bulletin 106, 135-59.

Craven L, Tuppen HA, Greggains GD, Harbottle SJ, Murphy JL, Cree LM, Murdoch AP, Chinnery PF, Taylor RW, Lightowlers RN *et al.* (2010) Pronuclear transfer in human embryos to prevent transmission of mitochondrial DNA disease. Nature 465, 82-5.

Dowling DK (2014) Evolutionary perspectives on the links between mitochondrial genotype and disease phenotype. Biochimica et biophysica acta 1840, 1393-403.

Ferreira CR, Burgstaller JP, Perecin F, Garcia JM, Chiaratti MR, Meo SC, Muller M, Smith LC, Meirelles FV, Steinborn R (2010) Pronounced segregation of donor mitochondria introduced by bovine ooplasmic transfer to the female germ-line. Biology of reproduction 82, 563-71.

Fu Q, Rudan P, Paabo S, Krause J (2012) Complete mitochondrial genomes reveal neolithic expansion into Europe. PloS one 7, e32473.

Hagelberg E (2003) Recombination or mutation rate heterogeneity? Implications for Mitochondrial Eve. Trends Genet 19, 84-90.

Hagstrom E, Freyer C, Battersby BJ, Stewart JB, Larsson NG (2014) No recombination of mtDNA after heteroplasmy for 50 generations in the mouse maternal germline. Nucleic Acids Res 42, 1111-6.

Hyslop LA, Blakeley P, Craven L, Richardson J, Fogarty NM, Fragouli E, Lamb M, Wamaitha SE, Prathalingam N, Zhang Q *et al.* (2016) Towards clinical application of pronuclear transfer to prevent mitochondrial DNA disease. Nature 534, 383-6.

Johnston IG (2015) Multiple hypothesis correction is vital and undermines reported mtDNA links to diseases including AIDS, cancer, and Huntingdon's. Mitochondrial DNA, 1-5.

Johnston IG, Burgstaller JP, Havlicek V, Kolbe T, Rulicke T, Brem G, Poulton J, Jones NS (2015) Stochastic modelling, Bayesian inference, and new in vivo measurements elucidate the debated mtDNA bottleneck mechanism. eLife 4.

Johnston IG, Williams BP (2016) Evolutionary Inference across Eukaryotes Identifies Specific Pressures Favoring Mitochondrial Gene Retention. Cell Systems 2, 101-11.

Koehler CM, Lindberg GL, Brown DR, Beitz DC, Freeman AE, Mayfield JE, Myers AM (1991) Replacement of bovine mitochondrial DNA by a sequence variant within one generation. Genetics 129, 247-55.

Latorre-Pellicer A, Moreno-Loshuertos R, Lechuga-Vieco AV, Sanchez-Cabo F, Torroja C, Acin-Perez R, Calvo E, Aix E, Gonzalez-Guerra A, Logan A *et al.* (2016) Mitochondrial and nuclear DNA matching shapes metabolism and healthy ageing. Nature.

Lippold S, Xu H, Ko A, Li M, Renaud G, Butthof A, Schroder R, Stoneking M (2014) Human paternal and maternal demographic histories: insights from high-resolution Y chromosome and mtDNA sequences. Investig Genet 5, 13.

Lott MT, Leipzig JN, Derbeneva O, Xie HM, Chalkia D, Sarmady M, Procaccio V, Wallace DC (2013) mtDNA Variation and Analysis Using Mitomap and Mitomaster. Current protocols in bioinformatics / editoral board, Andreas D Baxevanis [et al] 44, 1 23 1-6.

Mitalipov S, Amato P, Parry S, Falk MJ (2014) Limitations of preimplantation genetic diagnosis for mitochondrial DNA diseases. Cell reports 7, 935-7.

Monnot S, Gigarel N, Samuels DC, Burlet P, Hesters L, Frydman N, Frydman R, Kerbrat V, Funalot B, Martinovic J *et al.* (2011) Segregation of mtDNA throughout human embryofetal development: m.3243A>G as a model system. Human mutation 32, 116-25.

Morrow EH, Reinhardt K, Wolff JN, Dowling DK (2015) Risks inherent to mitochondrial replacement. EMBO reports 16, 541-4.

NCBI (2015) http://www.ncbi.nlm.nih.gov/nuccore/?term=%28015400%5BSLEN%5D:016700%5BSLEN%5D%29+AND+Homo%5BOrganism%5D+AND+mitochondrion%5BFILT%5D+NOT+%28Homo+sp.+Altai+OR+Denisova+hominin+OR+neanderthalensis+OR+heidelbergensis+OR+consensus+OR+ancient+human+remains%29.

Reinhardt K, Dowling DK, Morrow EH (2013) Medicine. Mitochondrial replacement, evolution, and the clinic. Science 341, 1345-6.

Sharpley MS, Marciniak C, Eckel-Mahan K, McManus M, Crimi M, Waymire K, Lin CS, Masubuchi S, Friend N, Koike M *et al.* (2012) Heteroplasmy of mouse mtDNA is genetically unstable and results in altered behavior and cognition. Cell 151, 333-43.

563 St John JC (2012) Transmission, inheritance and replication of mitochondrial DNA in
564 mammals: implications for reproductive processes and infertility. Cell and tissue
565 research 349, 795-808.
566 St John JC (2016) Mitochondrial DNA copy number and replication in reprogramming
567 and differentiation. Seminars in cell & developmental biology 52, 93-101.
568 St John JC, Facucho-Oliveira J, Jiang Y, Kelly R, Salah R (2010) Mitochondrial DNA
569 transmission, replication and inheritance: a journey from the gamete through the
570 embryo and into offspring and embryonic stem cells. Human reproduction update 16,
571 488-509.
572 Tachibana M, Sparman M, Sritanaudomchai H, Ma H, Clepper L, Woodward J, Li Y,
573 Ramsey C, Kolotushkina O, Mitalipov S (2009) Mitochondrial gene replacement in
574 primate offspring and embryonic stem cells. Nature 461, 367-72.
575 Takeda K, Tasai M, Iwamoto M, Akita T, Tagami T, Nirasawa K, Hanada H, Onishi A
576 (2006) Transmission of mitochondrial DNA in pigs and progeny derived from nuclear
577 transfer of Meishan pig fibroblast cells. Molecular reproduction and development 73,
578 306-12.
579 Taylor RW, Turnbull DM (2005) Mitochondrial DNA mutations in human disease.
580 Nature reviews Genetics 6, 389-402.
581 Treff NR, Campos J, Tao X, Levy B, Ferry KM, Scott RT, Jr. (2012) Blastocyst
582 preimplantation genetic diagnosis (PGD) of a mitochondrial DNA disorder. Fertility
583 and sterility 98, 1236-40.
584 Tsai T, St John JC (2016) The role of mitochondrial DNA copy number, variants, and
585 haplotypes in farm animal developmental outcome. Domestic animal endocrinology
586 56 Suppl, S133-46.
587 Wallace DC (2015) Mitochondrial DNA variation in human radiation and disease. Cell
588 163, 33-8.
589 Wallace DC, Chalkia D (2013) Mitochondrial DNA genetics and the heteroplasmy
590 conundrum in evolution and disease. Cold Spring Harbor perspectives in biology 5,
591 a021220.
592 Yamada M, Emmanuele V, Sanchez-Quintero MJ, Sun B, Lallos G, Paull D, Zimmer M,
593 Pagett S, Prosser RW, Sauer MV *et al.* (2016) Genetic Drift Can Compromise
594 Mitochondrial Replacement by Nuclear Transfer in Human Oocytes. Cell stem cell 18,
595 749-54.
596